Dramatic Structure and Social Status in Shakespeare's Plays

Heather Froehlich

ABSTRACT

This article discusses ways that dramatic structure can be analyzed through the use of social titles in Shakespeare's plays. Freytag's (1863) pyramid of dramatic structure is based on patterns he found in Shakespearean and Greek tragedy; more recently, computational methods are being employed to model narrative structure at scale. However, there has not yet been a study which discusses whether or not specific lexical items can be indicative of dramatic structure. Using Shakespeare's plays as an example, this essay fills the gap by observing how social titles can be used to explore the viability of narrative structure.

The nineteenth-century playwright, Gustav Freytag, argues that a clear division of action exists across the five acts from Shakespearean and Greek tragedy, based on repeated close readings of dramatic writing.[1] In this essay, I ask if this theory of dramatic structure can be tracked or otherwise measured using linguistic evidence. Focusing my attention on the linguistic class of vocatives, which include status terms (your majesty, madam, sir) to identify specific features of social identity, I hypothesize that the evocation of social class through the use of social referents paired with a name may regularly cluster around specific moments and aid in the construction of narrative, such as through act and scene divisions.[2] In this essay I test this assumption by observing simple frequencies at scale: I trace how the use of vocatives for social status are used in a bigram with a name occur in the whole of Shakespeare's plays and subsequently by genre (following conventions laid out by the First Folio).

This essay begins by describing Freytag's approach to dramatic structure, then discusses why the identification of class status can be a useful model for thinking about how to identify specific moments in narrative structure where status is explicitly mentioned. In section 3, I suggest that vocatives as ways of linguistically

identifying social status are a potentially salient piece of information for character construction. Moreover, it has potential to drive narrative action forward as a way to mark for recurrent characterization. Because social class is often encoded in vocatives directly referencing status or occupation, I argue that it is possible to assign class status based on vocative use and observe where in the narrative structure they are used. Social mobility underlies the storylines of many Shakespearean plays. Busse shows that nouns in the vocative mode foreground social structure in Early Modern English through semantically-derived relationships in Shakespeare's plays.[3] She illustrates vocative distributions by act in Shakespeare's plays, but assumes that vocative marking for social class is indifferent to dramatic structure.[4] She does not test this assumption, whereas I want to suggest that perhaps we can use vocatives as a way for re-establishing social identity throughout the narrative structure of a play-text.

Building on this, I produce a lexicon of potentially viable terms for status based on dramatis personae lists for Shakespeare's plays, then offer my own model of how to observe many social titles across narrative time using concordance plots. Concordance plots, a simple visualization showing locations where to find specific search terms in a linear visualization (from left to right, or 'start' to 'finish').[5] Freytag focuses his analysis in part on Shakespearean tragedies, no doubt due to Shakespeare's perceived centrality as the canonical early modern playwright. Based on Freytag's initial inquiry into tragedies, I extend my study to cover all of canonicized Shakespearean drama to observe if this theoretical framework scales to other genres. Moreover, as Shakespeare's plays are readily available in a variety of formats, it is possible to cross-reference status terms across more than one edition. To account for variation, I use two complementary digital editions of his plays as my source texts in Section 3: the Wordhoard implementation of the *Nameless Shakespeare* corpus and the Folger Digital Texts corpus.[6] To perform the concordance plot analysis in Section 4, I use the plain-text Shakespeare plays included in the *Shakespeare His Contemporaries* (Mueller 2015) corpus.[7] I replicate the concordance plot analysis four times: first, observing all of Shakespeare's plays in aggregate, then by individual genre (history-comedy-tragedy), following their classification by Mueller and Burns (2010) to observe if variation can be found across genre.[8] Finally, I offer some conclusions and routes for further study,

suggesting that vocatives are a productive metric for observing stylistic variation across genre.

Dramatic Structure

Freytag offers a well-received model of narrative structure. He uses Shakespearean and Greek tragedies as the basis for a pyramid shape for visualizing dramatic plots.[9] Freytag's model, commonly referred to as Freytag's Pyramid, was designed for the use of five-act plays.[10] In his description of narrative form, Freytag argues for a clearly visible division of action in dramatic texts based on repeated close reading of the texts. Figure 1 offers a visualization of Freytag's pyramid model.[11]
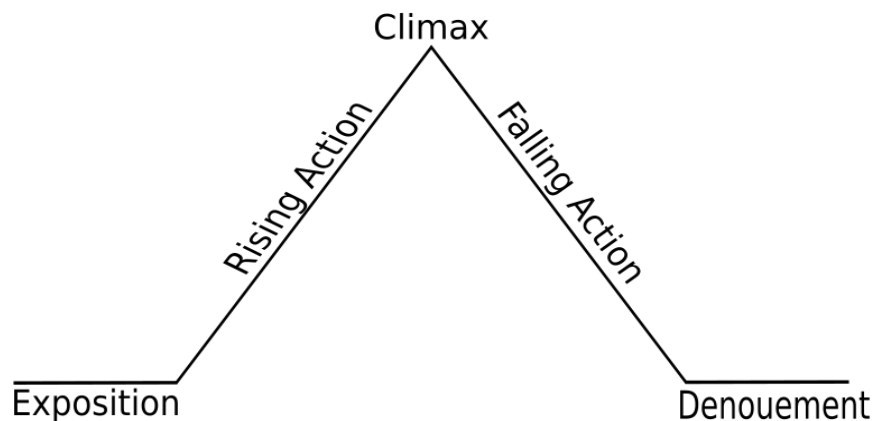

Figure 1. Freytag's pyramid, based on his study of five-act Shakespearean and Greek tragedies as it is commonly illustrated.

Freytag identifies a beginning with an "exposition", introducing the main characters and the situation at hand, and the "rising action", presenting a series of events which present the narrative's purpose. Frytag's pyramid offers a "climax" or high point of the narrative structure, with the language of the text moving towards a specific peak moment before a fall. This builds from the immediate introduction of the characters and plot at hand through the rising action into a climax. For Frytag specifically, this moment represents a turning point for the protagonist, in which their fate is changed irrevocably. Hamlet killing Polonius through the curtain in Gertrude's chamber is one such example of Frytag's climax, and the subsequent falling action, where Hamlet sets up the dumbshow to catch Claudius, serves to illustrate and foreshadow

Hamlet's downfall. The ending, which Freytag deems the "denouement", is a cathartic conclusion in which the plot is cleanly resolved in some way. The introduction of characters in the exposition would be a prime opportunity to establish distinguishing features such as identity markers. An emphasis on social class can be viewed primarily as a character-constructing event, rather than a non-vital piece of social information appearing sporadically throughout a linear dramatic text. Therefore, it is plausible to assume use of vocatives will cluster around specific moments in narrative structure. As Freytag suggests, dramatic narratives have broadly predictable structures: characters are introduced, something happens to them, and the end of the play must satisfactorily resolve the circumstances. To theorize this, we now move into a more socio-cognitive theoretical position.

In each defined moment of dramatic structure, the use of vocatives to identify characters by status can be hypothesized following Fiske and Neuberg's method for understanding and recognizing figures aims to direct attention at the construction of characterization and identification of figures within the available context.[12] Although Culpeper[13] notes that this model is not entirely aimed at literary characterization, it serves as a viable schema for conceptualization of new information and characters within the performed world of dramatic play-texts.[14] Moreover, characterization does not necessarily require understanding characters as real people: it is plausible to consider them through models of schemata, prototypicality, and/or social categories.[15] Regardless of the model used to consider new characters, they must be introduced and must be indicated through a variety of social cues, such as spoken language, dress, and address terms.

Following Fiske and Neuberg's theory of characterization, the use of vocatives in reference to a name should be greatest at the beginning of a play, introducing them following and then decrease until the end of the play, when it may be useful to be reminded of social status. After the introductory exposition, characters' social class will have been established and frequent reference should be unnecessary. By the time the climax has been reached, the vocative constructions mentioned in a bigram with a character's name should have decreased, as the act of constructing a characterization will have been cemented by now.

The characters with the most to lose are those with the highest social status; as a result these characters may require the highest amount of social maintenance. As one method to construct class is through overt references to social identity through the use of vocatives, a resurgence of character names in a bigram with a name may be anticipated at the peak of the dramatic structure.[16] These multimodal approaches converge to construct class during the falling actions. By the resolution of the play, social class should be re-established: social stratification and the potential for mobility reinforce the importance of social class. For example, a character changing position in the social hierarchy of Early Modern England through marriage might invoke a sudden shift in vocatives relating to social class; similarly, the impending downfall of a formerly powerful figure may invoke a reiteration of social class to emphasize the tragic figure's downfall. These would both be relevant moments to reiterate social identity through the use of vocative constructions.

Within this hypothesis, characters whose name is attached to a vocative in a bigram, such as *Lord Bigot* or *King Richard* or *Sir Roger* will be recurrent throughout the play, as whenever they are the referenced by name, their social title is included. While this helps reaffirm the social roles encoded in the texts overall, it does not actively hurt the process of characterization either – if anything, reiteration of the vocative serves as a reminder of characters. Additionally, the repetition of vocative titles may be a way to illustrate scene shifts, potentially as a way of reintroducing characters.

Social class in Early Modern England

Vocatives for social class often have a dual role as both a term of reference and a term of address, and the use of an honorific invokes a politeness rather than a strict naming practice; this only identifies the character, rather than the character's title.[17] This duality means that as politeness markers used for speaking to an individual above the speaker's social status, terms such as *lord* and *lady* also act as class status markers. These vocatives are indicative of social status for anyone, provided the addressee is higher in rank than the speaker. There are two dominant models of social class in Early Modern England. The first, offered by Nevalanien and Raumolin-Brunberg, introduces a model of social hierarchy for Early Modern England.[18] The classifications of social stratification outlined here by Nevalainen and Raumolin-Brunberg are initially designed for their Corpus of Early English Correspondence,

so they use literacy as their measure of social class. Their model accounts for gradiency from noblemen to gentlemen and professionals to non-gentry through honorific marking. Members of the gentry are especially literate and independent. This is presented in contrast to the illiterate non-gentry classes.

An alternative methodology for ranking characters by social status is modelled by the Sociopragmatic corpus.[19] Culpeper and Archer's class distinctions are determined by discoursal references rather than literacy, and requires much more manual annotation.

While Culpeper and Archer widen their scope by including trial transcripts, their classification scheme is heavily skewed towards the lower classes, who are more representative of the general population of Early Modern London. Busse also argues for analysing Shakespeare's use of vocatives in the clause through interpersonal co-reference, considering forms of semantic colouring through the use of n-grams as ways of informing Shakespearean vocative use and structure and uses patterns of Latinate adjectives or personal pronouns in collocate to colour vocatives functioning as an epithet.[20] In contrast, Nevalainen and Raumolin-Brunberg's social ranking schemas strive to be inclusive of multiple contexts in which one could be considered part of the gentry.

While Nevalainen and Raumolin-Brunberg emphasize differences between gentry and non-gentry in letter-writing, their approach allows for much more flexibility in identifying gentry figures for the needs of the study at hand.[21] They argue that the label "gentry" does not necessarily exclude the nobility and the monarchy in Early Modern England, whereas the label "commoners" explicitly excludes everyone else, accounting for a division between the commoner and non-commoner. For these reasons, this study follows Nevalaninen and Raumolin-Brunberg's decision to consider the non-commoners to be part of the gentry, offering a consistent division in social organization.

Identifying courteous titles

In this section I outline how to identify vocatives which specifically function as courteous titles. Importantly, Shakespeare's 36 plays are a manageable size for hand-

curation, and have the added benefit of being made widely available in many formats.[22] The dramatis personae, often supplied by editors, gives an independent rating of social status which I use to test my method of assigning status by vocative. To cover a range of opinions surrounding social class and nobility, I conflate information from two highly-annotated editions to get a wider view of paratextual information for Shakespeare's plays. For this task, I use two digitally-annotated editions of Shakespearean drama based on well-established critical print editions: the Wordhoard Shakespeare corpus and the Folger Digital Texts corpus.[23] Crucially, the Wordhoard and Folger Digital Text editions of Shakespeare's plays contain slightly different information about each character in their dramatis personae lists, which provides the possibility to develop a more nuanced view of a character status.[24]

One such example can be found in *1 Henry VI*. Henry Beaufort is described by the Folger Digital Texts as "bishop of Winchester and afterwards cardinal". While this makes it clear that Henry Beaufort is a high-status individual in *1 Henry VI*, WordHoard includes that he is also the great-uncle to the King. This phenomenon provides additional context which only one edition may not show. Using information from the Folger Digital Texts cross-referenced with the Wordhoard corpus, one can determine the minimum number of high-status figures in each Shakespeare play just through dramatis personae lists. Dramatis personae lists tend to highlight speaking roles, and these digitally-encoded texts will include moments where all or both characters on stage will speak. It is difficult to discern the total number of characters in each play, as an indefinite number of non-speaking parts could be available.[25] Due to the intrinsic difficulty in identifying the total number of characters in a play, a minimum quantity of upper-class figures by gender will be presented.

Cross-referencing social identity from two editions is especially useful when considering how to categorise minor characters. For example, Wordhoard gives the following information for *The Comedy of Errors* in the Dramatis Personae:

> Solinus, duke of Ephesus
> Aegeon, a merchant of Syracuse
> Twin brothers, and sons to Aegeon and Aemilia
>   Antipholus of Ephesus

   Antipholus of Syracuse
 Twin brothers, and attendants on the two Antipholuses
   Dromio of Ephesus
   Dromio of Syracuse
 Balthazar, a merchant
 Angelo, a goldsmith
 First Merchant, friend to Antipholus of Syracuse
 Second Merchant, to whom Angelo is a debtor
 Pinch, a schoolmaster
 Aemilia, wife to Aegeon an abbess at Ephesus
 Adriana, wife to Antipholus of Ephesus
 Luciana, her sister
 Luce, servant to Adriana
 A Courtezan
 An Officer
 A Servant
 A Gaoler
 And other Attendants

Among the named characters (Balthazar, Dromio of Ephesus, Dromio of Syracuse, Pinch) and several unnamed but unique characters ("a courtesan", "An Officer", "a Gaoler", "First Merchant") there are also an unknown number of Attendants. Although Wordhoard often only encodes the characters with speaking roles, the Folger Digital Texts use a numbering convention of MinorCharacter X to list each unnamed minor character with a speaking role in its underlying XML. Each minor character with a speaking role gets an individual number in the Folger Digital Texts (e.g. MinorCharacter 1, MinorCharacter 2) but the total number of minor characters remains unclear. Not all of the minor characters have speaking roles, but could be listed as an unknown quantity of 'lords'. In other words, the total number of non-speaking minor characters for each play remains debatable.

For example, characters with royal or noble titles included in their listing in the dramatis personae, such as 'Ferdinand, king of Navarre' (*Love's Labours Lost*) would immediately be considered part of the upper classes, but not the gentry. Similarly, characters who fall under a heading, such as 'lords attending on the King: Berowne, Longaville, Dumain' (*Love's Labours Lost*) would all be considered upper class as well. In addition, the following rules were applied, iteratively:

1. If related to a figure of the upper classes (e.g. married, child of) still part of the upper classes
2. Servants are not necessarily members of the gentry class
3. Unnamed multiple lords do not contribute to the final counts
4. Officers and soldiers are gentry figures, following Nevalainen and Raumolin-Brunberg's Table 4.1, which places them largely in the gentry rather than the non-gentry.
5. Merchants are gentry, following Nevalainen and Ramoulin-Brumberg's Table 4.1, Model 3.

Rule 3 in particular accounts for minor characters who are unnamed, unspecified, and listed in dramatis personae, especially minor characters who be considered nobles or gentry by virtue of their title, but the total number of minor characters remains unclear. For this reason, the quantitative results produce a view of 'At least X percent of the dramatis personae for each play represents characters from the upper classes'. There is another confounding issue underpinning this analysis: there are also various members of the court such as attendants and servants who are not explicitly members of the gentry.

Citizens, neighbours and various members of the larger community are also quite unlikely to be gentry, as they represent the common man and woman, especially in history plays. There are also characters such as murderers, witches, soothsayers and messengers who have also not been counted as members of the higher status, as their social rank is unclear from their included characterization. Each character must be specifically a part of the upper classes through the invoking of rank either explicitly (through titles and/or status) or implicitly (through relationships which directly link figures to gentry status). Compiling a list of characters – who may not have sufficient details of their social status included in the dramatis personae – to the vocative title and name bigrams provides us with the ability to infer social class based on status. Although ambiguous and/or polysemous titles such as *mistress* may not be as straightforward as creating a rule that all characters with the title of duke are members of the nobility or that all named lord characters are part of the gentry, the identification of vocatives means it is possible to categorize characters by status in Shakespeare's plays using their title paired with a name.

Moreover, each title investigated here correlates exactly to the social status of an individual character. The vocative title *king*, for example, is used only for heads of state and *earl* and *duke* are applied to figures of that status. It would be rare to find evidence of up-talking through the use of higher-status titles applied to a character of lower status: titles which reference the nobility are not going to be applied to anyone who is not in that specific social class, due to the rigidity of Early Modern English social structures. As a result, class-specific titles applied characters are almost always applied to named male characters of that specific social class. So while *sir*, *lord* and *master* can be used interchangeably towards any non-commoner male character, the titles *esquire/squire*, *duke* and *earl* are more closely tied to a specific social status and identity.

Meanwhile, the social titles available for women in Shakespeare's plays are more complicated. For example, the title *lady* describes a woman with authority over servants, attendants and serves as a head of the household in addition to functioning as an honorific for lower-status servants. *Lady* is "a term referring to a woman to indicate social superiority, as the head of a household or one who rules over subjects, servants or attendants. The title was honorific, referring to a woman of high rank" but "paradoxically, it was also used to signify a servant to a woman of high rank, such as those who waited on a queen or noblewoman and might well be nobly born themselves"[26]. The title of *lady* is therefore also applicable to any high-status character who also serves another high-status character. *Mistress* is another complex social title: it alternately denotes a female head of a household or a woman with authority over servants and attendants.[27] But the title *mistress* can also be a term for prostitutes, and *miss* is a diminutive form of *mistress*, which usually appears with a name and "is a form of polite address to a married woman, or an unmarried woman or girl".[28] So not all characters that take the title "Mistress" are part of the gentry-class, although the term suggests that the character in question could be. For example, Mistress Quickly from *Merry Wives* and *Henry IV parts 1* and *2* is a professional, rather than a member of the gentry class, whereas Mistress Bianca (from *Taming of the Shrew*) would be a gentry figure.

Understandably, there is some distrust surrounding vocatives as a reliable way to identify social class because this process requires heavy contextualization within the dramatic world of the play-text. For example, "'Sir' in Sir Credulous gives us a

reasonable identification that he is of gentry status but 'sir' alone used by one individual to address another provides no reliable indication of status".[29] That said, it is possible to use regular expressions to pair each relevant title with a name using regular expressions where every example of the search string will be identified in a bigram with a proper noun. Regular expressions are string-matching queries producing matches on a symbol-by-symbol basis which can be then used to identify specific strings which are more complex than a simple letter-matching query.

By using the search string *queen [A-Z]*, every instance of letters which matches the pattern of *queenspaceAnyCapitalLetter* produces entries such as *queen Elizabeth, queen Dido, queen Elinor, queen Mab, queen Bess, queen Margaret, queen Mary,* among others. The search strings used in the regular expressions are case insensitive, so proper-noun *Queen* or all-lowercase *queen* left-adjacent to a string which contains a capital letter after the space right-adjacent to the *n* are both identifiable. This process produces a list of every instance from the corpus through a list of concordance lines containing every instance of the specified conditions and provides the filename the line was pulled from. However, this process will occasionally produce irrelevant hits, such as instances of the first-person singular pronoun *I* as right-adjacent to the title query, producing irrelevant outputs but accurate to the query such as the string *queen I.* While examples of *title I* are indeed within the parameters of the query provided, they do not produce evidence of a title paired with a name and are discarded. With the aid of a text editor, this query can be double-checked to ensure relevant examples, such as *queen Isabel*, would be retained.

Using a concordance program supporting regular expressions, such as AntConc, a search string which matches the general pattern of *title [A-Z]* can be used to identify all proper nouns which are referenced through the use of a vocative for social status in a bigram in a corpus.[30] This process produces a list of every instance from the corpus through a keyword in context viewer.[31] Based on the titles referenced in the dramatis personae lists, and supplemented by other potentially synonymous forms with the aid of the Historical Thesaurus of the Oxford English Dictionary[32], the below table lists all potential titles under investigation.[33]

sir [A-Z]	lady [A-Z]
master [A-Z]	mistress [A-Z]
lord [A-Z]	mrs [A-Z]
king [A-Z]	goodwife [A-Z]
duke [A-Z]	goody [A-Z]
earl [A-Z]	widow [A-Z]
goodman [A-Z]	madam [A-Z]
sirrah [A-Z]	miss [A-Z]
signor [A-Z]	missus [A-Z]
signior [A-Z]	dame [A-Z]
father [A-Z]	queen [A-Z]
friar [A-Z]	gentlewoman [A-Z]
squire [A-Z]	countess [A-Z]
esquire [A-Z]	
marquis [A-Z]	

Table 1. All regular expressions included in the search queries from

These titles have been arranged in order of social status and by gender. Although this is not designed to implicitly skew male, there are two more titles for men than women.[34] Once irrelevant examples, such as *title I* are manually removed from the results, it is possible to use simple frequency counting to observe which vocatives for social class Shakespeare uses in his writing. Table 2 shows that although these terms are potentially available for use, Shakespeare does not use all of them.

Vocative title as a regular expression	Total relevant examples
sir [A-Z]	399
master [A-Z]	144
lord [A-Z]	303
king [A-Z]	150
duke [A-Z]	33
earl [A-Z]	3
goodman [A-Z]	2
sirrah [A-Z]	7
signor [A-Z]	0
signior [A-Z]	102
father [A-Z]	21
friar [A-Z]	61
squire [A-Z]	0
esquire [A-Z]	0
brother [A-Z]	84

marquis [A-Z]	5
lady [A-Z]	66
mistress [A-Z]	111
mrs [A-Z]	0
goodwife [A-Z]	0
goody [A-Z]	0
widow [A-Z]	6
madam [A-Z]	12
miss [A-Z]	0
missus [A-Z]	0
dame [A-Z]	0
queen [A-Z]	14
gentlewoman [A-Z]	0
countess [A-Z]	2

Table 2. Vocative titles in use in Shakespeare's plays

There are more mentions of status titles which mark for male characters than those for female characters, and some titles are very rarely used. *Squire* is not used at all in Shakespeare's corpus, although squires are contemporary to the period. It is possible that they may be considered amongst the unnamed characters, or that their title was never mentioned: after all, boys who aspire to be knights may not necessarily be titled, either. The clergy covers a variety of roles, including bishops, deacons, priests, pastors and ministers; bishops are established members of the gentry, whereas the role of a deacon is a less prestigious title than priesthood and ministry. Table 6 also considers *friar* and *father* as potential gentry titles, as clergy are often on the cusp of gentryhood, following Nevalainen and Raumolin-Brunberg.[35] Though priests and fathers are more likely to be members of the gentry class because they are ordained, Table 6 shows that Friar is more frequent than Father, probably due to the presence of Friar Laurence in *Romeo and Juliet.*

The noted a lack of *signor* attached to a name in the Shakespeare corpus is very worrisome, as Italian gentlemen are common roles in comedies such as *Taming of the Shrew, Merchant of Venice,* and *Much Ado about Nothing.* These characters are referenced using the Italian title of address for men repeatedly in the plays, but none of these results come up in a regular expression search for *signor [A-Z]*, because they are considered under the variant spelling *signior*. It is quite likely that in preprocessing the texts, *signor* was corrected to become *signior*. Rather than making

the dominant spelling *signor* the default, it has become regularized to a variant spelling.[36] Because this is a fairly frequent politeness marker for gentlemen in plays set in Italy, it is simple enough to substitute a spelling variation producing results into the list of terms for further investigation.

And although female characters are present in Shakespearean drama, Table 2 shows that they are less likely to be addressed by title and their name. The one exception, *mistress*, can function both as a social title and as part of a naming convention. A character can be called *mistress* as an honorific and as part of her character's name in a way unavailable to other titles.[37] Although *lady* is less likely to be used to describe a high-status female character as part of a name, Lady Macbeth is the most prominent character using this construction in Shakespeare's plays. Instead, *mistress* appears to be the title of choice, and is used for characters such as Mistress Quickly and Mistress Ford. Mistress may be used in place of *missus* and *miss*, which also may contribute to these terms' overall absence. Other female-gendered forms are generally absent in comparison to the male-gendered social titles.

With a sense of which terms are more frequent than others in Shakespeare, following the lists of vocatives outlined in Tables 1 and 2, I now focus my attention on the highest-frequency vocatives from Table 6. Because examples with $n > 20$ are so infrequent, their contributions to an overall picture may introduce noise to an aggregated view. To observe overall use of these terms in the plays, one can construct a regular expression which consists of the most-frequent forms in from Table 2. Titles chosen are strung together by pipelines, so that the full query asks for 'or' rather than 'and'. The full search string reads as follows:

> lord [A-Z]|sir [A-Z]|master [A-Z]|duke [A-Z]|earl [A-Z]|king [A-Z]|signior [A-Z]|lady [A-Z]|mistress [A-Z]|madam [A-Z]|queen [A-Z]|dame [A-Z]

When the query encounters a play without one or more of these specific regular expressions, it will continue to loop through until all parts in the query have run. With this search string, it is now possible to observe how the most-frequently appearing bigrams appear in Shakespeare's plays, and whether or not their

invocation has any correlation to the idea of dramatic structure with the aid of a simple visualization process.

Applying models of dramatic structure to Early Modern drama

Where Freytag uses a graph to model the overall appearance of dramatic structure, I am more interested in observing the placement of specific lexical items throughout a text. Applying AntConc's concordance plot visualization[38] to the Shakespeare plays included in Muller's 2014 *Shakespeare His Contemporaries* corpus is a simple way to view lexical frequency in a text, normalized for length. Like Freytag's visualizations of dramatic structure, they are read from left ("start") to right ("finish"). A dark line represents one instance of the search query in use; white space means the search query is not present. Converting from a keyword-in-context concordance view to a concordance plot view, each hit is translated into one line; each play gets its own concordance plot, again to be read from left to right (or start to finish). Any character string matching the above regular expressions are counted and recorded as one hit.[39] Several plays are provided as an example in Figure 1 below.
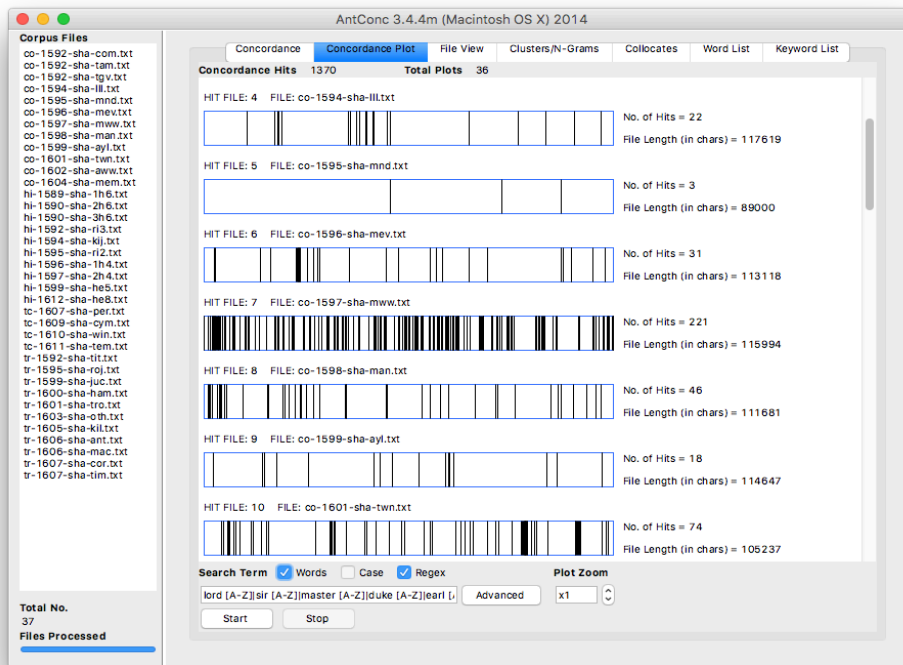

Figure 2: Concordance plots as seen in the AntConc interface

Individually, these plays do not cover a lot of examples. *Merry Wives of Windsor*, labelled as 'mww' above (third concordance plot from the top and from the bottom), has quite a lot of examples of our search string, whereas *As You Like It* ('ayl', second from bottom) shows comparatively little usage. But in aggregate they represent a much larger part of dramatic language. Once each individual concordance plot from Figure 1 is saved as an individual image, it is possible to combine each concordance plot into an aggregated, average view of all concordance plots for an overall picture of vocative use across the aggregated corpus.

Despite the popularity of the so-called 'bag of words' model which ignores lexicogrammar and syntax[40], this essay follows the theoretical and statistical position voiced by Kilgarriff (2005) and echoed by Evert (2005), Gries (2005) and Dunning (1993). He correctly notes that the use of language can be neither random nor arbitrary, and in fact governed by a range of syntactic rules, making the null hypothesis that words can appear more often than simply by chance impossible:

> Language is non-random and hence, when we look at linguistic phenomena in corpora, the null hypothesis will never be true. Moreover, where there is enough data, we shall (almost) always be able to establish that it is not true. In corpus studies, we frequently do have enough data, so the fact that a relation between two phenomena is demonstrably nonrandom, does not support the inference that it is not arbitrary. Hypothesis testing is rarely useful for distinguishing associated from non-associated pairs of phenomena in large corpora. Where used, it has often led to unhelpful or misleading results.[41]

Lijffijt et al agree: "the use of the X^2 and log-likelihood ratio tests is problematic in this context, as they are based on the assumption that all samples are statistically independent of each other."[42] While Lijffijt et al do argue later in their article that there are moments when such a test would be appropriate, such as the study of variation and statistical significance between discrete corpora,[43] this study is not looking at totally discrete corpora but rather subcorpora culled from one coherent body of work broadly conceived as belonging to a solo author.[44] Therefore, the present essay adheres to Kilgarriff's initial claim, and is bolstered by the fact that literary language is even less likely to be randomly distributed, making a chi-test or

other significance score especially inappropriate. Following Zipf's law[45], literary language is even less likely to be randomly distributed: aside from the question of lexicogrammar, overall vocabulary frequency is governed by type: the most frequent terms in any given corpus are high-frequency function words, whereas content-driven language is far less frequent throughout the corpus and must be distributed in a way that presents a meaningful narrative. For these reasons, this project foregoes questions of statistical significance for the rhetoric of visualization.

I use Forster's methods[46] for aggregating images using ImageMagick, a command-line software package for image manipulation, makes it possible to produce average concordance plots for all of Shakespeare's plays to see if vocatives marking for social class are visible at the level of narrative structure.[47] Forster describes this process as "taking 50 images setting each to 2% transparency and then stacking them all together so that you have a single page […] it is comparing the same pixel location across multiple images, so size is key".[48]

Despite the differences in each play's individual word counts, each concordance plot is identical in dimensions regardless of length, meaning it is possible to consider each play as part of a larger aggregated whole. Following Forster's advice, each concordance plot was cropped to an identical size and layered to construct an aggregated image.[49] Figure 3 shows vocatives for social class in Shakespeare's plays can appear anywhere:

Figure 3. Average view of 36 concordance plots for Shakespeare's plays

While there are moments of relative darkness (presence) and lightness (absence), there are no sustained moments where a vocative for social class is completely absent. Even if one is to divide Figure 3 into thirds representing the beginning, middle and end of a play, these vocatives are equally available throughout the narrative structure. The fact that these plays use vocatives to mark for social titles should not be a surprise. Figure 4 shows that there is no coherent, sustained gap in the whole of the corpus. However, the question remains if this will hold when we

consider plays based on generic classification. Does genre have any influence in how characters' social status is presented?

The larger the corpus, the more information which has to get layered into the average concordance plot. This may be a limiting factor in Figure 3. To reduce some of this noise, I split the corpus into three discrete genres following the First Folio conventions: 14 comedies, 11 histories and 11 tragedies.[50] Applying the same methodology as outlined above, a generic study of Shakespeare's plays will allow us to see if his use of vocatives for social class is consistent across gender. I will begin with Shakespeare's history plays, then move to his comedies and end with his tragedies.

History plays include many members of court, making them especially heavily skewed towards gentry. Many of the examples are embedded in character names, so it is not tremendously surprising to find that the Shakespeare's histories use vocatives for social class extremely frequently. However, while there are lots of examples available in the history plays, based on characterization alone, Figure 4 shows that there is no particular sustained portion of these plays which are without vocatives.

Figure 4. Average view of Shakespeare's histories

Like Figure 3, which offers the overall picture of Shakespeare's plays, Figure 4 shows that vocatives can appear anywhere in the history plays. As a genre, history plays are quite stable for literary scholars; they centre around historical figures with established status and social ranks. Thus it is should not be a huge surprise that there is no general pattern of sustained absence in these plays. Figure 5 simply confirms what literary scholars have long known about this particular genre, which is that it is primarily concerned with retelling stories of historical figures of rank and status.

Meanwhile, it could be argued that comedy plays are just as much about social class as histories: a female character's change in status from daughter to wife is the widely

recognized defining feature of a comedy. This shift in status for a female character would be very closely tied to her impending marriage; the narrative of the Shakespeare's comedies centres around the success of this change in social status. This makes comedies intrinsically unlike histories, as the plot functions quite differently. As a result, it may be expected that the first third may establish individuals' social status with a heavy use of vocatives. The shift in status surrounding a successful match would introduce a second rise in vocative use in the final third of the play, with a more marked decrease in the middle third. However, this proposed outcome is quickly proven untrue in Figure 5.

Figure 5. Average view of Shakespeare's comedies

Figure 5 illustrates that vocatives for social class again do not show any particular patterns or clusters in the start, middle, or end of plays. There is no large, clearly defined moment where there is a noticeable absence of vocatives, nor is there a clear moment where vocatives are most strongly concentrated in Figure 3, Figure 5, or Figure 4. Thus far, none of vocatives under investigation do offer any strong correlation to a five-act dramatic structure.

Given all the evidence presented thus far, it is reasonable to expect that Shakespeare's tragedies will continue to show this kind of trend. But, as Figure 6 below shows, Shakespeare's tragedies show a different distribution than the previous two genres. Figure 5 shows a marked decrease in vocatives for social class around the middle half of the aggregated concordance plot, and a slight resurgence near the middle of the final third, but again a noticeable absence at the end of the final third. Unlike in histories and comedies, where these vocatives could appear anywhere in the narrative structure, Shakespeare's tragedies include comparative periods of absence.
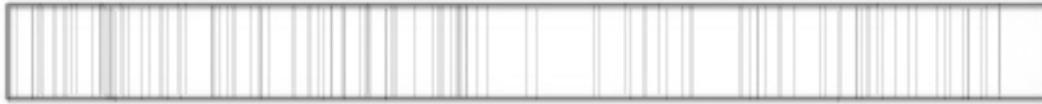
Figure 6. Average view of Shakespeare's tragedies

Distribution is not nearly as uniform here as in Figure 3, 4 or 5; Figure 6 also shows that the tragedies have fewer examples of our vocatives overall.[51] This is especially noticeable around the halfway point in the middle third of the overall image, which shows a much-decreased frequency of vocatives for social class to the point of showing extended stretches with extremely infrequent use (marked by white spaces). This is in stark contrast to Figures 4 and 5, where the terms under investigation are heavily used throughout the plays. Instead, the use of vocatives for social class in direct reference to a characters' name completely drops off by the end of the play and are broadly missing from the middle third.

Figure 6 therefore may be indicative of Freytag's theory of a literary climax, in which he claims that Shakespearean and Greek tragedies reach their emotional high point somewhere in the third act in a five-act structure. By broadly refraining from using vocatives for social class at this particular moment, Shakespeare's tragedies may indeed illustrate a specific kind of shift in style in which linguistic politeness strategies and social class become briefly less important. The sudden lack of vocatives marking for social class between male and female characters suggests that the breakdown of the dialectic begins around Act III. The tragic downfall of the heroic Self could be marked by a temporary shift in style and tone away from conventional courteously terms during this time.[52]

As Figure 6 suggests, the social status of the main character is motioned to rather than made explicit during the construction of the tragic Self's downfall. This gap is available for each Shakespearean tragedy around roughly the same general location in the play as figure 7 illustrates. However, the exact moment is slightly different in each play, as Table 3 outlines:

Play name	Part of play without vocatives for social class as a naming strategy
Titus Andronicus	From II.xx to IV.i
Romeo and Juliet	From II.iv to III.iii

20

Julius Caesar	Lord Brutus: from II.i to IV.iii; master Antony appears once in III.i.
Hamlet	From III.i to IV.ii
King Lear	Two instances: once in IV.iii and IV.v
Troilus	From III.ii to IV.i
Othello	From II.iii to V.i
Antony and Cleopatra	From III.vi to IV.v
Timon of Athens	From III.iv to IV.i

Table 3. Points in Shakespearean tragedies without vocatives for social class attached to a name

Macbeth and *Coriolanus* have no instances of these particular vocatives, so they are excluded from Table 3. A lack of vocatives marking for social class in a bigram with a name suggests Act III broadly marks the beginning of tragic hero's downfall and foreshadows the impending tragic ending. They reappear around the start of Act IV in each play. In *Hamlet*, during the gap between instances of *lord Hamlet* (Act III, scene I, and Act IV, scene 2), Hamlet tells Horatio to watch Claudius during their dumbshow. The players frame King Claudius, and Hamlet readies himself behind a curtain to kill the praying King Claudius; Gertrude demands to talk to her son and he instead murders Polonius, and finally in Act IV scene 2 Gertrude reflects on Lord Hamlet's actions, and his title of "Lord Hamlet" returns. After setting actions in motion implicating Claudius and between his decision to enact revenge and his rash change of heart to kill the man behind the curtain, Hamlet has set his downfall into motion. A similar example is available in *Titus Andronicus*, where Act III as a whole is devoid of vocatives marking social class with a name, though phrases marking for status are used throughout Titus' investigation of Aaron and Luicius' abuse towards Lavinia. Titus wants information out of them, and he must appease them to get the answers he wants: he must use face-saving strategies to coerce them into confessing their actions. The success of his efforts is questionable, leading to Titus' decision to take revenge after Lavinia's ravishing.

In *Antony and Cleopatra*, this phenomenon is delayed slightly, occurring between III.6 and IV.5, but the sudden absence of vocatives marking for an individual character by status also occurs; here, Antony escapes from Rome to Egypt to fight alongside Cleopatra. The parts of the play without reference to naming practices including vocatives for social class are during Antony's battle, and he blames Cleopatra for making him abandon his previously noble self in his pursuit for her affection. Finally in this section of the play, Caesar prepares to conquer Antony's

army. While Cleopatra shoulders the blame for Antony's defeat, his role as tragic hero is cemented at a great cost to Cleopatra. In each example outlined here, the tragic downfall of a character is contingent on others responding to their attempts at face-saving strategies and getting information out of them as to why they acted as they did, and introduces the tragic hero's downfall from Act III onwards.

The visual reintroduction of vocatives as part of a naming strategy we see in Act IV suggests that this relative absence in the middle of tragedies may be an additional distinguishing feature of Shakespearean tragedies from the rest of his dramatic corpus. Instead of using explicit naming, Shakespearean tragedies offer a highly localized form of social titles; rather than emphasizing one individual's status as tragic hero, we see him using politeness markers and register-raising strategies such as "sir", "my lord", "your honour", and others towards others. He aims to flatter those around him in hopes of getting them to act on his behalf and/or to encourage them to divulge information to him. Of course, this backfires, and prompts the tragic hero to enact a series of events leading to his downfall. The shift back into vocatives as a naming strategy around Act IV and into the beginning of Act V serves primarily as reminders of the tragic hero's status within the world of the play. This is very much in contrast to Shakespeare's comedies and histories, which show frequent use from start to finish in the concordance plots.

The primary finding of this study is Shakespeare's tragedies are sufficiently different to the comedies and histories in their use of vocatives marking explicitly for class. This suggests several new directions for the study of genre and vocatives. In particular, there may be something specific about tragedy that affects ways that social status terms are used. First, it would be good to replicate this study on both a larger selection of authors and a larger scale more generally. This may be indicative of something that makes Shakespearean tragedy different from other playwrights in Early Modern England writing tragedies, or it may be a larger generic difference which can be traced. With that in mind, I would also be very reluctant to call this a universal feature without ample research across many periods and many years, which will require a much larger lexicon of status terms. In addition, this may be something which Freytag had initially identified in his initial claims about tragedy in his study of dramatic structure. Because Freytag's model is primarily based on Shakespearean and Greek tragedy, another route would be to replicate this study

using the corpora driving other studies and updating the terms for investigation accordingly to see if it their models can account for other kinds of vocative features. This study also suggests that simple visualizations are more productive than complex graphs for observing lots of lexical examples at once.

Conclusions

In this essay, I investigate ways that courteous or otherwise socially elevated titles appear as part of character name, to show that the dramatis personae lists from Shakespeare's plays are a useful way to categorise individuals by social status. I show that issues of socially courteous titles are completely independent of the process of characterization and narrative. In Shakespearean drama, vocatives marking for social class are identifiable through the use of concordance plots. Following Freytag's visualization of dramatic structure, this essay shows that concordance plots are an alternative model of visualizing narrative structure. This study shows that there is no particularly strong correlation between vocatives for social class and dramatic structure except in Shakespeare's tragedies. Their noted absence offers several new research questions for further investigation about genre in plays written by Shakespeare and his contemporaries.

As the overall picture of Shakespeare's plays (Figure 4) shows, the use of vocatives to mark out individual characters does not appear to have any bearing on character or narrative construction. Figure 4 shows that vocatives marking for social class appear equally throughout the whole concordance plot rather than a strong pattern of presence and absence at any particular moment. It is possible what Figure 4 shows is an aggregated view of scene shifts in many plays at once, where characters move on and off stage and their social title becomes a salient feature for identifying the character.

In performance, the divisions of acts and scenes are often imperceptible; act and scene divisions may well be an artefact of print culture rather than a feature of the plays in the first instance. In Shakespeare's plays, scene length is far from normalised; very short scenes of a hundred lines or less can be juxtaposed with comparatively longer scenes. However, it is difficult to trace this, as each concordance plot is normalized for length regardless of how many words are in each

text. Furthermore, scene length is also quite irregular; there is no standard scene length. And, very short scenes of a hundred lines or less can be juxtaposed with comparatively longer. As a result, scenes do not necessarily correspond; moreover, imposing even a broad act division model on this image may be retrofitting a structure to the texts which is not necessarily accurate. With this in mind, I would also like to suggest that each instance of a vocative is not necessarily indicative of a specific character's presence on stage. Although there are individual moments without reference to vocatives for social class, it would be very difficult to identify exactly where these absences would be precisely located and if they overlap completely: although these concordance plots have been normalized for length, it is very difficult to identify exact moments across a layered image like this. In performance, the divisions of acts and scenes are often imperceptible, and act and scene divisions may well be an artefact of print culture rather than a feature of the plays in the first instance, and that the many repeated instances of social titles referencing a character may be indicative of scene changes or other forms of movement across the corpus. This study therefore also suggests that markers for social class may be not be always be visible at the level of dramatic structure.

I would therefore like to close by offering one final possibility about the visualizations offered. The uses of vocatives for names serve as a way to re-introduce and re-establish characters entering and exiting the stage. Here, social re-identification becomes interactional, allowing individuals to be re-established alongside other social cues such as costuming and setting and also enacting motive-driven narrative techniques. A performative strategy to imply closeness and/or intimacy between two or more character, this is an effective way of manipulating someone into completing the speaker's desired effect, such as acting performing courtship practices, acting on behalf of another, or exposing the hearer of the speaker's intentions.

Each instance of a vocative is therefore not necessarily indicative of a specific character's presence on stage; instances may be more indicative of characters following social courtesy based on rank.[53] Each time socially courteous titles are used, it does not necessarily mean that they are on stage, or that they are even part of the action at hand. Referencing these characters by their full name with the title as a feature of social politeness and modes of social courtesy in the period means

that they can be used regardless of whether or not the character in question is on-stage. As Early Modern stage productions required actors to play several roles, this may be a way to re-establish and constantly reappraise character's status throughout the plays rather than limiting descriptions of characters early in the plays.[54] Therefore, when we visualize language at scale like this this, we are really observing very small interactive moments which aggregate into a larger narrative whole, which is more complicated than practitioners of quantitative methods have anticipated.

Notes

[1] Freytag, Gustav, and Elias J. MacEwan, trans. *Freytag's Technique of the Drama: An Exposition of Dramatic Composition and Art. An Authorized Translation from the 6th German Ed. by Elias J. MacEwan*. (1900 [1863]).

[2] Huddleston, Rodney D., and Pullum, Geoffrey K. *The Cambridge Grammar of the English Language*. (2002) 522-523; Busse, Beatrix, *Vocative Constructions in the Language of Shakespeare*. (2006).

[3] Busse, Beatrix. *Vocative Constructions in the Language of Shakespeare*. (2006): 133; 137.

[4] Busse, Beatrix.. *Vocative Constructions in the Language of Shakespeare*. (2006): 77-79; 91-92; 415-423.

[5] Anthony, L. AntConc (Version 3.3.5) [Computer Software]. Tokyo, Japan: Waseda University. http://www.antlab.sci.waseda.ac.jp/ (2013).

[6] Martin, P. Burns, et al. *Wordhoard,* Version 1.4.4,. Web applet. (1 March 2011) http://wordhoard.northwestern.edu/userman/index.html; Mueller, Martin.. "The Nameless Shakespeare". *TEXT Technology: the journal of computer text processing* 14.1. (2005).; Folger Shakespeare Library. *Shakespeare's Plays* from *Folger Digital Texts*. Ed. Barbara Mowat, Paul Werstine, Michael Poston, and Rebecca Niles. (2014.) http://folgerdigitaltexts.org/.

[7] The exact texts used in this study came from an early version of the *Shakespeare His Contemporaries* corpus, which was published during the writing of this essay (Mueller 2015). See the associated Dataverse repository to replicate this study.

[8] This essay does not take contemporary debates about authorship into consideration, instead using Wordhoard's divisions of genres as its guiding principle.

[9] Freytag, Gustav, and Elias J. MacEwan, trans. *Freytag's Technique of the Drama: An Exposition of Dramatic Composition and Art. An Authorized Translation from the 6th German Ed. by Elias J. MacEwan*. (1900 [1863]).

[10] A similar structure is outlined by Aristotle in *Poetics,* but it has been argued that this is better suited to three-act plays. Freytag's pyramid has been the dominant mode of understanding dramatic structure since 1863, and lifts heavily from Aristotle's three modes of protasis, epitasis and catastrophe (18.1-2). While an Aristotelian model of dramatic structure in Early Modern drama seems to have fallen out of fashion by the 1980s, Freytag's model continues to be an accepted narrative structure by literary scholars.

[12] Fiske, S. T., & Neuberg, S. L. A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, Vol. 23. (1990): pp. 4-8.

[13] Culpeper, Jonathan. *Language and Characterisation: People in Plays and Other Texts*. (2001). 85.

[14] Literary studies in particular are indebted to Barthes (1970 [1974]), in which lists of character traits are considered; Frye (1957: 171) prefers a pragmatic model of characterization, in which presentation influences character.

[15] Bartlett, F.C.. *Remembering: A Study in Experimental and Social Psychology.* (1995 [1932]; Rosch, E. et al. "Basic Objects in Natural Categories". *Cognitive Psychology* 8.3 (1976) 382–439; Propp, Propp, V. *Morphology of the Folktale: Second Edition*. Austin: University of (2010 [1968]).

[16] Alternate modes of constructing social class, which may be used in collaboration with the use of vocatives include (re)establishing setting, include emphasizing clothing (e.g. "enter Lear, fantastically dressed with weeds", which stresses how far he has removed himself from courtly dress).

[17] Demmen, Jane Elizabeth Judson. "Charmed and Chattering Tongues: Investigating the Functions and Effects of Key Word Clusters in the Dialogue of Shakespeare's Female Characters." (2009) 146.

[18] Nevalainen, Terttu, and Helena Raumolin-Brunberg. 'Sociolinguistics and Language History: The Helsinki Corpus of Early English Correspondence'. *Journal of Linguistics* 13. (1994). p.140, based on Laslett, Peter. *The World We Have Lost* (1983): 38.

[19] Archer, Dawn, and Jonathan Culpeper. 'Sociopragmatic Annotation: New Directions and Possibilities in Historical Corpus Linguistics'. *Corpus Linguistics by the Lune: A Festschrift for Geoffrey Leech*. Ed. Paul Rayson, Anthony McEnery, and Andrew Wilson, Geoffrey N. Leech. (2003): 37–58; Archer, Dawn. *Questions and Answers in the English Courtroom (1640–1760): A Sociopragmatic Analysis*. (2005); *Sociopragmatic Corpus.* A Derivative of *A Corpus of English Dialogues 1560–1760*, Compiled Under the Supervision of Merja Kytö (Uppsala University) and Jonathan Culpeper (Lancaster University), (2007); Lutzky, Ursula. *Discourse Markers in Early Modern English* (2012).

[20] Busse, Beatrix. *Vocative Constructions in the Language of Shakespeare*. (2006): 109-127, 211-231

[21] For example, Palander-Collin (2007) discusses the need for multiple sociopragmatic annotation schemes, arguing that different projects have different needs, such as conversational analysis compared to variationist sociolinguistics (online, section 2).

[22] This manual curation model is very unlikely to scale well, especially for dramatic plays that is less well-documented than the Shakespearean corpus. However, for an individual author's corpus or a small corpus comprising a sample of several different authors, it would certainly be possible to apply the methods displayed here.

[23] Martin, P. Burns, et al. *Wordhoard,* Version 1.4.4,. Web applet. (1 March 2011) http://wordhoard.northwestern.edu/userman/index.html; Mueller, Martin.. "The Nameless Shakespeare". *TEXT Technology: the journal of computer text processing* 14.1. (2005).; Folger Shakespeare Library. *Shakespeare's Plays* from *Folger Digital Texts*. Ed. Barbara Mowat, Paul Werstine, Michael Poston, and Rebecca Niles.

[24] Dramatis personae lists are usually considered paratextual and often editorial in nature. This is part of the reason it is worth looking at more than one specific edition of the plays, though theoretically this could be extended to many more editions for a more robust view.

[25] In addition, instances of speech assigned to 'all' characters or 'both' characters (if there are just two) have been removed from this analysis.

[26] Findlay, Alison. *Women in Shakespeare: A Dictionary*. (2010): 226

[27] Findlay, Alison. *Women in Shakespeare: A Dictionary*. (2010): 271

[28] Findlay, Alison. *Women in Shakespeare: A Dictionary*. (2010): 271. Moreover, the diminutive form *miss* does not appear in the Shakespeare corpus. Perhaps the diminuitive was less popular in the early modern period, whereas *mistress* was more widely used, though this requires further study with a larger corpus.

[29] Archer, Dawn, and Jonathan Culpeper. 'Sociopragmatic Annotation: New Directions and Possibilities in Historical Corpus Linguistics'. *Corpus Linguistics by the Lune: A Festschrift for Geoffrey Leech*. (2003): 53

[30] Anthony, L. AntConc (Version 3.3.5) [Computer Software]. Tokyo, Japan: Waseda University. http://www.antlab.sci.waseda.ac.jp/ (2013).

[31] The search strings used in the regular expressions explored in Table 4 are case insensitive to account for variation, so that capitalization for the social title is irrelevant, but the word to its immediate left must be capitalized.

[32] See Froehlich (2017).

[33] Kay, Christian, Jane Roberts, Michael Samuels, and Irené Wotherspoon (eds.). *The Historical Thesaurus of English*, version 4.2. Glasgow: University of Glasgow. (2015) http://historicalthesaurus.arts.gla.ac.uk/; Froehlich, Heather. *Social Identity in Shakespeare's Plays: A Quantitative Study.* 2017

[34] As literary critics such as Phyllis Rackin (2000, 44) claim, "plays with overtly repressive and misogynist themes […] are held up as historically accurate expressions of beliefs generally held in Shakespeare's time." The majority of writing from Early Modern England is by men and social history suggests that men were more involved in matters of court. However, a lack of courteous titles for women does not necessarily claim that women could not be the head of state or exert powerful influence over Early Modern London. For example, prior to her marriage to Henry VIII, Anne Boleyn had an enormous influence at court; her daughter, Queen Elizabeth I, ruled the country from 1558 to 1603.

[35] Nevalainen, Terttu, and Helena Raumolin-Brunberg. *Sociolinguistics and Language History: Studies Based on the Corpus of Early English Correspondence*. (1996): 58.

[36] *Signior* is included as a spelling variation according to the OED. (see 'Signor, N.' *OED Online*. *Oxford English Dictionary*. Web. 9 Jan. 2015.)

[37] Findlay, Alison. *Women in Shakespeare: A Dictionary*. (2010): 271.

[38] Anthony, L. AntConc (Version 3.3.5) [Computer Software]. Tokyo, Japan: Waseda University. http://www.antlab.sci.waseda.ac.jp/ (2013).

[39] Where previous discussions of vocative *title [A-Z]* included omitting examples of *title [first person pronoun I]* here they have been retained, which skews the data somewhat: vocative *title I* functions far more as a politeness marker than a character-naming mechanism. This can conflate two functions of the vocatives.

[40] Lijffijt, Jefrey, Terttu Nevalainen, Tanja Säily, Panagiotis Papapetrou, Kai Puolamäki, and Heikki Mannila. "Significance Testing of Word Frequencies in Corpora". *Digital Scholarship in the Humanities*. (Advanced access, 8 December 2014): 4; Lijffijt, Jefrey, Terttu Nevalainen, Tanja Säily, Panagiotis Papapetrou, Kai Puolamäki, and Heikki Mannila. "Significance Testing of Word Frequencies in Corpora". *Digital Scholarship in the Humanities*. (Advanced access, 8 December 2014): 2.

[41] Kilgarriff, Adam. "Language Is Never, Ever, Ever, Random". *Corpus Linguistics & Linguistic Theory* (November 2005): 263–76. https://doi.org/10.1515/cllt.2005.1.2.263. pp. 272-273. See also Gries, S. Th. "Null-hypothesis significance testing of word frequencies: a follow-up on Kilgarriff". *Corpus Linguistics and Linguistic Theory* 1, no. 2 (November 2005): 277–94; Evert, S. "The Statistics of Word Cooccurrences: Word Pairs and Collocations". Dissertation, Institut für Maschinelle Sprachverarbeitung, University of Stuttgart (2005); and Dunning, T. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics* (1993). Vol.19: 61–74.

[42] Lijffijt, Jefrey, Terttu Nevalainen, Tanja Säily, Panagiotis Papapetrou, Kai Puolamäki, and Heikki Mannila. "Significance Testing of Word Frequencies in Corpora". *Digital Scholarship in the Humanities*. (Advanced access, 8 December 2014: 1; George K. Zipf. *The Psychobiology of Language*. Houghton-Mifflin (1935); George K. Zipf. *Human Behavior and the Principle of Least Effort*. Cambridge, Massachusetts: Addison-Wesley (1949).

[43] Lijffijt, Jefrey, Terttu Nevalainen, Tanja Säily, Panagiotis Papapetrou, Kai Puolamäki, and Heikki Mannila. "Significance Testing of Word Frequencies in Corpora". *Digital Scholarship in the Humanities*. (Advanced access, 8 December 2014: 4).

[44] Authorship for Shakespearean plays is complicated – but generally scholars follow the 1623 first folio (see Jowett, John. Shakespeare and Text. (2007); Smith, Emma. *Shakespeare's First Folio: Four Centuries of an Iconic Book* (2016) and Smith, Emma. *The Cambridge Companion to Shakespeare's First Folio* (2016)) for generic classification. The present project follows this practice.

[45] George K. Zipf. *The Psychobiology of Language*. Houghton-Mifflin (1935); George K. Zipf. *Human Behavior and the Principle of Least Effort*. Cambridge, Massachusetts: Addison-Wesley (1949).

[46] Forster, Chris. "Vellum in War Time: Playing with MJP data". *Chris Forster.* Web. http://cforster.com/2011/11/playing-with-mods/ 2011.

[47] ImageMagick is freely available from http://www.imagemagick.org/. I am deeply indebted to Chris Forster for his proof-of-concept using magazine pages (2011, http://cforster.com/2011/11/playing-with-mods/); more technical detail is available in the following blog posts, Froehlich 2015a and b: http://hfroehli.ch/2015/02/24/how-to-address-many-concordance-plots-at-once-2/ and http://hfroehli.ch/2015/02/26/a-cautionary-tale/)

[48] Forster, Chris. "Vellum in War Time: Playing with MJP data". *Chris Forster.* Web. http://cforster.com/2011/11/playing-with-mods/ 2011.

[49] This is not a composite image but an averaged image. ImageMagick has an option for creating images using the command 'composite', which requires specifying opacity, much like in Adobe Photoshop's layer function. In theory these are both commands which produce aggregated views of many images together. The 'composite' command requires setting thresholds, whereas the 'average' command operates on many images at a default 2% opacity.

[50] The Folio classification for genre is a widely accepted convention in the field.

[51] The tragedy subcorpus covers the same number of plays covered by the comedy subcorpus, so this cannot be an effect of size. Although they are not equal in wordcount, the visualization is normalized for size, so that each play – regardless of length – is visualized across the same space.

[52] Following Bamber, Linda. *Comic Women, Tragic Men: Study of Gender and Genre in Shakespeare*. (1982): 48.

[53] cf Nevalainen, Terttu, and Helena Raumolin-Brunberg. "Sociolinguistics and Language History: The Helsinki Corpus of Early English Correspondence." *Hermes, Journal of Linguistics* 13 (1994): 135–14; Nevalainen, Terttu,

and Helena Raumolin-Brunberg. 1996. *Sociolinguistics and Language History: Studies Based on the Corpus of Early English Correspondence* (1996).

[54] Please see Ichikawa, Mariko. *The Shakepearean Stage Space.* Cambridge UP: Cambridge, England. (2013) and Gurr, Andrew and Mariko Ichiwaka. *Staging in Shakespeare's Theatres.* Oxford Shakespeare Topics: Oxford, England. (2000) for discussions of staging in Shakespeare's plays – including the crucial interventions of entrances and exits, especially for roles played by more than one person.